# Quantitative Mass Spectrometric Multiple Reaction Monitoring Assays for Major Plasma Proteins\*

## Leigh Anderson‡§ and Christie L. Hunter¶

Quantitative LC-MS/MS assays were designed for tryptic peptides representing 53 high and medium abundance proteins in human plasma using a multiplexed multiple reaction monitoring (MRM) approach. Of these, 47 produced acceptable quantitative data, demonstrating within-run coefficients of variation (CVs) (n = 10) of 2–22% (78% of assays had CV <10%). A number of peptides gave CVs in the range 2-7% in five experiments (10 replicate runs each) continuously measuring 137 MRMs, demonstrating the precision achievable in complex digests. Depletion of six high abundance proteins by immunosubtraction significantly improved CVs compared with whole plasma, but analytes could be detected in both sample types. Replicate digest and depletion/digest runs yielded correlation coefficients (R<sup>2</sup>) of 0.995 and 0.989, respectively. Absolute analyte specificity for each peptide was demonstrated using MRM-triggered MS/MS scans. Reliable detection of L-selectin (measured at 0.67  $\mu$ g/ml) indicates that proteins down to the  $\mu$ g/ml level can be quantitated in plasma with minimal sample preparation, yielding a dynamic range of 4.5 orders of magnitude in a single experiment. Peptide MRM measurements in plasma digests thus provide a rapid and specific assay platform for biomarker validation, one that can be extended to lower abundance proteins by enrichment of specific target peptides (stable isotope standards and capture by anti-peptide antibodies (SISCAPA)). Molecular & Cellular Proteomics 5:573-588, 2006.

Accurate quantitation of proteins and peptides in plasma and serum is a challenging problem because of the complexity and extreme dynamic range that characterize these samples (1). The widely adopted separative (survey) approach to proteomics, in which an attempt is made to detect all components, has proven to be limited in sensitivity toward low abundance proteins (2) and typically provides limited quantitative precision. The alternative candidate-based approach, which relies on specific assays optimized for quantitative detection of selected proteins, can provide significantly in-

From the ‡The Plasma Proteome Institute, Washington, D. C. 20009-3450 and ¶Applied Biosystems, Foster City, California 94404 Received, October 7, 2005, and in revised form, November 28,

2005 Published, MCP Papers in Press, December 6, 2005, DOI 10.1074/ mcp.M500331-MCP200 creased sensitivity (into the pg/ml range) and precision (CVs<sup>1</sup> < 5-10%) at the cost of restricting discovery potential toward novel proteins (3). In practice, a combination of these approaches (one or more survey approaches for *de novo* biomarker discovery coupled with a candidate-based approach to biomarker validation in large sample sets) appears to provide an effective staged pipeline (4) for generation of valid plasma biomarkers of disease, risk, and therapeutic response.

Candidate-based specific assays rely on the specificity of capture or detection methods to select a specific molecule as analyte. Capture reagents such as antibodies can provide extreme specificity (particularly when two different antibodies are used as in a sandwich immunoassay) and form the basis of most existing clinical protein assays. There is intense interest in miniaturizing sets of such assays (5, 6) in array formats (on planar substrates, beads, etc.), although significant problems remain in the production of suitable antibodies and in the simultaneous optimization of multiple assays in one fluid volume.

Mass spectrometry provides an alternative assay approach, relying on the discriminating power of mass analyzers to select a specific analyte and on ion current measurements for quantitation. In the field of analytical chemistry, many small molecule analytes (e.g. drug metabolites (7), hormones (8), protein degradation products (9), and pesticides (10)) are routinely measured using this approach at high throughput with great precision (CV < 5%). Most such assays use electrospray ionization followed by two stages of mass selection: a first stage (MS1) selecting the mass of the intact analyte (parent ion) and, after fragmentation of the parent by collision with gas atoms, a second stage (MS2) selecting a specific fragment of the parent, collectively generating a selected reaction monitoring (plural MRM) assay. The two mass filters produce a very specific and sensitive response for the selected analyte that can be used to detect and integrate a peak in a simple one-dimensional chromatographic separation of the sample. In principle, this MS-based approach can provide

<sup>&</sup>lt;sup>1</sup> The abbreviations used are: CV, coefficient of variation; QqQ, triple quadrupole; MRM, multiple reaction monitoring; S/N, signal-to-noise; Nat, natural sample-derived peptide; SIS, stable isotope-labeled internal standard; polySIS, polyprotein SIS; SISCAPA, stable isotope standards and capture by anti-peptide antibodies; MIDAS, multiple reaction monitoring-initiated detection and sequencing; MARS, multiple affinity removal system.

absolute structural specificity for the analyte, and in combination with appropriate stable isotope-labeled internal standards (SISs), it can provide absolute quantitation of analyte concentration. These measurements have been multiplexed to provide 30 or more specific assays in one run (11). Such methods are slowly gaining acceptance in the clinical laboratory for the routine measurement of endogenous metabolites (e.g. in screening newborns for a panel of inborn errors of metabolism (12)) and some drugs (e.g. immunosuppressants (13)).

Recently the MRM assay approach has been applied to the measurement of specific peptides in complex mixtures such as tryptic digests of plasma (3). In this case, a specific tryptic peptide can be selected as a stoichiometric representative of the protein from which it is cleaved and quantitated against a spiked internal standard (a synthetic stable isotope-labeled peptide (14)) to yield a measure of protein concentration. In principle, such an assay requires only knowledge of the masses of the selected peptide and its fragment ions and an ability to make the stable isotope-labeled version. C-reactive protein (3), apolipoprotein A-I (15), human growth hormone (16), and prostate-specific antigen (17) have been measured in plasma or serum using this approach. Because the sensitivity of these assays is limited by mass spectrometer dynamic range and by the capacity and resolution of the assisting chromatography separation(s), hybrid methods have also been developed coupling MRM assays with enrichment of proteins by immunodepletion and size exclusion chromatography (18) or enrichment of peptides by antibody capture (SISCAPA (19)). In essence the latter approach uses the mass spectrometer as a "second antibody" that has absolute structural specificity. SISCAPA has been shown to extend the sensitivity of a peptide assay by at least 2 orders of magnitude (19) and with further development appears capable of extending the MRM method to cover the full known dynamic range of plasma (i.e. to the pg/ml level).

There is compelling evidence that high and medium abundance plasma proteins have value as clinical biomarkers and thus that there may be applications for specific MRM assays even without antibody enrichment. One may therefore ask how many plasma proteins can be measured by quantitating their peptides in a plasma digest, and how precise could these measurements be? If the measurement strategy proves to be robust, could it be carried out using existing high throughput LC-MS/MS platforms? To address these questions we generated and tested MRM assays based on peptides from a variety of high to medium abundance plasma proteins to see how many could be measured effectively by LC-MS/MS with and without subtraction of the most abundant proteins. An understanding of the performance of MS/MS in this application could enable routine and relatively inexpensive measurement of classical plasma proteins and also provide a foundation for use of MS/MS in more sensitive anti-peptide antibody-enhanced SIS-CAPA assays for low abundance proteins.

### EXPERIMENTAL PROCEDURES

Peptide Selection and MRM Design—We developed a set of MRMs iteratively using three basic approaches: pure *in silico* design from sequence databases, design from available LC-MS/MS proteomic survey data, and comprehensive MRM testing of all of the candidate peptides of a protein.

In our initial attempt to generate MRMs by purely in silico methods, a set of 177 proteins and protein forms was assembled that are demonstrated or have potential to be plasma markers of some aspect of cardiovascular disease (20), and a subset of 62 proteins was selected for which an estimate of normal plasma abundance was available. Predicted tryptic peptides for each of these were generated along with relevant Swiss-Prot annotations and a series of computed physicochemical parameters: e.g. amino acid composition, peptide mass, Hopp-Woods hydrophilicity (21), and predicted retention time in reversed-phase (C18) chromatography (22). An index of the likelihood of experimental detection was derived from a data set reported by Adkins et al. (23) by counting the number of separate "hits" for the peptide in the data set divided by the number of hits for the most frequently detected peptide from the same protein. An overall index of peptide quality was generated according to a formula that gave positive weights to Pro, KP, RP, and DP content and negative weights to Cys, Trp, Met, chymotrypsin sites, certain Swiss-Prot features (carbohydrate attachment, modified residues, sequence conflicts, or genetic variants), and mass less than 800 or greater than 2000. The 3619 tryptic peptides predicted for the 62 protein marker candidates (6-497 peptides per target) ranged in length from 1 to 285 amino acids. Within the useful range of 8-24 amino acids, 721 peptides had a C-terminal Lys and 690 had a C-terminal Arg. In this report, peptides from 30 of these target proteins ending in C-terminal Lys were selected for further study.

We also selected peptides based on two types of direct proteomic survey experiments. In the first case we carried out classical LC-MS/MS analysis of plasma digests in which the major ions observed were subjected to MS/MS using the ion trap capabilities of the 4000 Q TRAP instrument. The identified peptides showing the best signal intensity and chromatographic peak shape for a given parent protein were selected. In addition, we used the Global Proteome Machine database of Craig *et al.* (24) to select peptides from target proteins that were frequently detected (multiple experiments).

Finally we used an adaptation of the MIDAS workflow, described previously for discovering post-translational modifications (25, 26), to look for measurable tryptic peptides from a variety of plasma proteins. In this approach, the protein sequence is digested *in silico*, likely y-ion fragments are predicted, and theoretical MRMs are generated for all the peptides in an acceptable size window. These MRMs are then used as a survey scan in a data-dependent experiment to detect specific peptide peaks, and each resulting MRM peak is examined by full scan MS/MS to obtain sequence verification of the hypothesized peptide. To verify peptide specificity in designated protein targets, selected peptides were searched with BLASTP for exact matches against the genome-derived human, mouse, and rat Ensembl peptides using Ensembl Blastview (www.ensembl.org/Homo\_sapiens/blastview).

*"Random" MRMs*—Two approaches were used to generate pseudorandom MRMs. In the first case we used 100 MS1 values distributed randomly (by the Excel RAND function) between 408.5 and 1290.2 (the maximum and minimum of an early set of real MRMs we tested) paired with MS2 values chosen randomly between this MS1 and the maximum of the real MRMs (1495.6), thus mimicking the properties of our real MRMs (which are generally +2 charge state peptides and +1 charge fragments). In a second set we paired 131 MS1 values chosen randomly from among MS1 values in a large table of real MRMs with MS2 values chosen randomly from the real MS2 values of the same list, imposing only the constraint that each MS2

had to be between 1 and 2 times the paired MS1 mass (to approximate our selection criteria for real MRMs). Peaks detected in these MRMs were examined by triggering MS/MS (the MIDAS workflow).

Reagents—The following chemicals were used: trypsin (Promega), sodium dodecyl sulfate (Bio-Rad), iodoacetamide (Sigma), formic acid (Sigma), tris-(2-carboxyethyl)phosphine (Sigma), and acetonitrile (Burdick and Jackson).

Plasma Depletion and Digestion—All experiments were performed on aliquots of a single human plasma sample from a normal volunteer. The six highest abundance proteins were depleted from plasma using the multiple affinity removal system ("MARS"; Agilent Technologies spin columns) according to the manufacturer's protocol. Depleted sample was then exchanged into 50 mM ammonium bicarbonate using a VivaSpin concentrator (5000 molecular weight cutoff, Viva-Science). Undepleted plasma was also desalted before digestion.

Both depleted or undepleted plasma samples were denatured and reduced by incubating proteins in 0.05% SDS and 5 mm tris-(2-carboxyethyl)phosphine at 60 °C for 15 min. The sample was then adjusted to 10 mm with iodoacetamide and incubated for 15 min at 25 °C in the dark. Trypsin was added in one aliquot (protease:protein ratio of 1:20) and incubated for 5 h at 37 °C.

Labeled Peptide Internal Standards: polySIS-A series of SIS peptides was added to samples in selected experiments by spiking with the tryptic digest of a "polySIS" polyprotein.<sup>2</sup> Briefly this protein was produced by cell-free transcription and translation of a synthetic gene coding for 30 concatenated tryptic peptide sequences (derived from 30 plasma proteins) in the presence of U-13C6,U-15N2-labeled lysine (a total mass increment of 8 amu compared with the natural peptide). The 30 peptides were selected based on our initial in silico MRM design approaches and are thus not fully optimized using experimental data. Of these peptides, 13 were used in the present studies (the remainder were not reproducibly detected in digested plasma with peak area >1E+04). The positioning of the label atoms at the extreme C terminus of each peptide has the effect that all fragments that contain the C terminus (i.e. the y-ions) will show the mass shift due to the label, whereas all the fragments that contain the N terminus (and hence have lost one of more C-terminal residues: the b-series ions) will have the same masses as the corresponding fragments from the natural (sample-derived) target protein. These features (shifted y-ions and normal b-ions) provide a simplification in interpreting the fragmentation patterns of the SIS peptides in comparison with the similar QCAT concept described recently by Beynon et al. (27). To determine the absolute concentration of polySIS protein, an aliquot was diluted with 1 M urea, 0.05% SDS, and 50 mM Tris, pH 8 and subjected to N-terminal Edman sequencing, yielding an initial concentration of 5  $\pm$ 1 pmol/µl. A tryptic digest of the polySIS protein was spiked into whole and depleted human plasma digests at the final concentrations shown in Table I.

*LC-MS/MS Analysis*—Plasma digests with and without added polySIS peptides were analyzed by electrospray LC-MS/MS using LC Packings (a division of Dionex, Synnyvale, CA) or Eksigent nanoflow LC systems (Table I) with 75- $\mu$ m-diameter C<sub>18</sub> PepMap reversedphase columns (LC Packings) and eluted with gradients of 3–30% acetonitrile with 0.1% formic acid. A column oven (Keystone Scientific, Inc.) was used to maintain the column temperature at 35 °C. Electrospray MS data were collected using the NanoSpray® source on a 4000 Q TRAP hybrid triple quadrupole/linear ion trap instrument (Applied Biosystems/MDS Sciex), and the peaks were integrated using quantitation procedures in the Analyst software 1.4.1 (Intelli-Quan algorithm). MRM transitions were acquired at unit resolution in both the Q1 and Q3 quadrupoles to maximize specificity.

### RESULTS

Design of MRM Assays for Abundant Plasma Proteins-In an initial approach to the selection of representative peptides for MRM assays, a single peptide of 8-18 amino acids was chosen from each of 30 proteins spanning a broad range of plasma concentrations (6.6  $\times$  10<sup>8</sup> down to 1 fmol/ml normal concentration) based on computed characteristics alone (19). MRMs were designed assuming doubly charged peptide ions and using fragments selected as likely y-ions above the m/z of the 2+ parent ion with collision energies assigned by a generic formula (CE =  $0.05 \times m/z + 5$ ), and the peptides were expressed as a concatamer polySIS protein containing single copies of each peptide labeled with [U-13C6,U-15N2]lysine. When a tryptic digest of the polySIS was analyzed, all 30 peptides were detected by MRMs. When a digest of whole human plasma was added to the polySIS peptides, 19 of the labeled polySIS peptides were still detected by the same MRMs, but only 11 of the plasma digest-derived unlabeled cognate peptides were detected (by the same MRMs adjusted for isotope label masses).

Because different peptides from a single protein can vary widely in detectability by ESI-MS, we attempted to improve upon the in silico approach to MRM design using experimental data from a conventional peptide survey scan of a human plasma digest and applying the selection criteria to peptides with demonstrated detectability. Using a 3-h LC gradient, MS/MS scans were collected for the major doubly or triply charged ions across the separation using information-dependent data acquisition, and a second run was performed using time-filtered exclusions of the peptide ions detected in the first run. The combined results identified 54 plasma proteins ranging in abundance from albumin down to fibronectin (normal plasma concentration of  $\sim$ 300  $\mu$ g/ml). This experimental MS/MS data provided explicit information for peptide selection, charge state, and most abundant y-ion m/z value under the specific conditions used (i.e. electrospray ionization with collisional peptide fragmentation), allowing improved design of MRMs. When these MRMs were then used to analyze the same sample in a subsequent run, triggering MS/MS scans at any MRM signal, most of the peptides were detected as peaks in the chromatogram and identified by database search as expected. In most of these MRM chromatograms, only a single peak was detected.

Because peptide detection sensitivity using MRM is expected to be greater than that achieved in a full scan MS survey approach, a comprehensive *de novo* MRM design method was explored for those proteins not detected in the above survey experiment. Using a novel software tool, a large set of MRMs was generated for each of a series of target proteins by selecting all predicted tryptic peptides in a useful size range together with multiple high mass y-ion fragments of each (the "MIDAS" workflow (25, 26)). These MRMs were then tested in LC-MS/MS runs of the unfractionated plasma digest,

<sup>&</sup>lt;sup>2</sup> L. Anderson and C. L. Hunter, manuscript in preparation

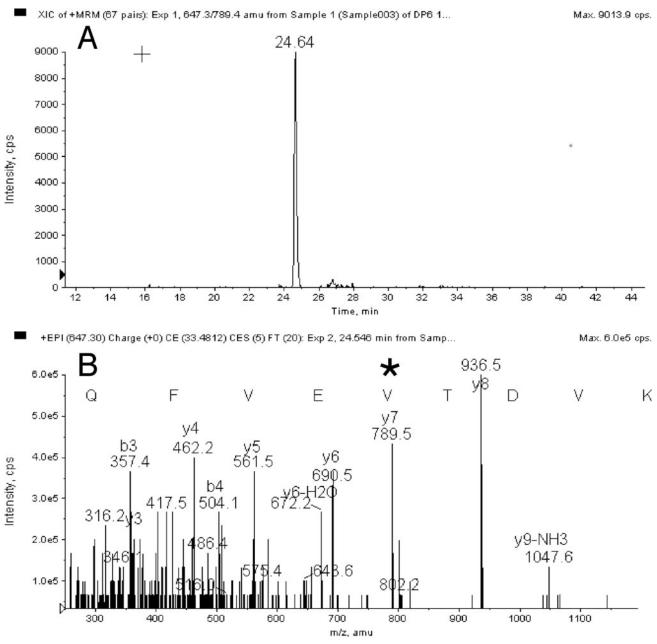


Fig. 1. MIDAS workflow (MRM-triggered MS/MS) verification of the identity of peptide DLQFVEVTDVK representing fibronectin (normal concentration,  $\sim$ 300 µg/ml) in a digest of depleted human plasma. *A* shows the ion current profile of MRM 647.3/789.4 (MS1/MS2), and *B* shows the MS/MS spectrum of peptide fragments taken at the time of the peak (\* marks the y7 fragment monitored in this MRM). *cps*, counts per second.

grouped in panels that included all the predicted tryptic peptides of one or two proteins at a time (50–100 MRMs per run), with MS/MS scans triggered on any peaks observed. Of 12 proteins examined, nine produced at least one usable MRM (signal-to-noise (S/N) ratio >20).

MRM results from the above approaches were pooled, and a set of optimized MRMs was assembled that covered a total of 60 peptides representing 53 proteins in human plasma (Table II; seven proteins were represented by two peptides). This set includes 18 peptides selected by the *in silico* approach (indicated by an X in the SIS column of Table II): eight of the initial 30 *in silico* peptides were eliminated as likely to be too low abundance for detection, and better alternative peptides were selected from experimental data for four others. For all but one of the peptides we elected to measure two fragments (*i.e.* using two MRMs per peptide), yielding 119 MRMs. Finally we included MRMs for 18 stable isotope-labeled internal standard ("SIS") versions of target peptides (*i.e.* the tryptic digest of the polySIS protein) spiked into the digest plasma samples. The resulting set of 137 MRMs was meas-

TABLE I Summary design of data sets (experiments A–F)

Replicate runs were performed in series with 30-min washes between runs. Load is expressed as the equivalent volume of plasma from which the sample was derived. Load factors express total or non-depleted (MARS flow-through) loads relative to experiment A.

<b>F</b>	Replicate	Comple		Equivalent plasma	l	PolySIS		
Experiment	runs	Sample	LC system	volume loaded	Total protein	Non-depleted proteins	spike	
				μΙ			fmol	
А	10	Depleted plasma digest	LC Packings	0.01	1	1	1.3	
В	10	Whole plasma digest	LC Packings	0.01	10	1	1.3	
С	10	Whole plasma digest	Eksigent	0.001	6	0.1	2.0	
D	10	Depleted plasma digest	Eksigent	0.01	0.6	1	2.0	
E	10	Depleted plasma digest	Eksigent	0.033	3.3	3.3	6.0	
F1_1	4	Depletion 1, digest 1	Eksigent	0.01	1	1		
F1_2	4	Depletion 1, digest 2	Eksigent	0.01	1	1		
F2_1	4	Depletion 2, digest 1	Eksigent	0.01	1	1		
F2_2	4	Depletion 2, digest 2	Eksigent	0.01	1	1		

ured in all the replicate runs described below using an 18-ms dwell time per MRM and a resulting cycle time of  ${\sim}3$  s between measurements.

After the final MRM method was constructed, each MRM transition and respective retention time were validated again as indicative of each specific peptide. Full scan MS/MS was acquired upon the appearance of the MRM signal, and each resultant spectrum was manually inspected to determine matching to the specific peptide (Fig. 1).

Application to Digests of Whole and Depleted Plasma-Six experiments (A-F) were performed in each of which the same set of 137 MRMs was measured during sequential replicate LC-MS/MS runs of a single sample (same injection volume), and the appropriate peaks were integrated by the Analyst software to yield a value (peak area) for each MRM in each run. Experiments A-E (10 replicate runs each) are summarized in Table I. These experiments included tryptic digests of both whole (unfractionated) human plasma (experiments B and C) and plasma depleted of abundant proteins by MARS immunoaffinity subtraction (experiments A, D, and E), high (experiments B and E) and low (experiments A, C, and D) total peptide loadings, and different chromatographic setups. Our objective was to assess the performance of the MRMs in various typical plasma digest experiments. The reference peptide load (experiment A) was derived from tryptic digestion of the protein contained in 10 nl of plasma ( ${\sim}700~\text{ng}$  of total protein assuming an average plasma concentration of ~70 mg/ml (28)) after depletion of the most abundant proteins using an Agilent MARS spin column (~84% of protein mass should be removed based on a calculation using normal abundances). This loading, comprising an estimated 110 ng of total peptides, proved to be a loading compatible with very good nanoflow chromatography of the MRM peptides. Experiments B and E used higher loadings to explore the trade-off between peak stability (chromatographic quality, adversely affected by increased load) and S/N ratio (improved by increased analyte quantity). We concluded that the 110 ng loading was optimal. Chromatographic elution times were quite reproducible, showing average CVs of 2% (experiment D) and 2.5% (experiment E).

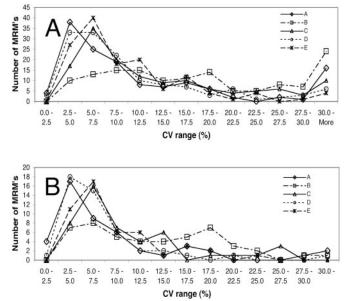


FIG. 2. Histogram of CVs of MRM values (peak areas) for five experiments (experiments A–E) across all 137 MRMs (A) or across the 47 "best" MRMs (B).

Fig. 2A shows histograms of the coefficients of variation for the five experiments (individual values for each MRM are contained in Table II). In analyses of depleted plasma, more than 60% of the MRMs show within-run CVs less than 10%, and almost half have CVs below 5%. A number of these MRMs (e.g.  $\alpha_1$ -antichymotrypsin, apolipoprotein E, hemopexin, heparin cofactor II, plasminogen, prothrombin, fibrinogen  $\gamma$  chain, complement C4, and factor B) showed an average within-run CV of 3-4% across three experiments, precision equivalent to that of good clinical immunoassays. Analyses of whole (undepleted) plasma digests showed generally higher CVs (20-50% of MRMs had CV <10%). It is important to note that these reproducibility measures are computed on raw peak areas without correction using internal standards. Four of the measured proteins were expected to be removed by the depletion process used (Agilent MARS TABLE II

A set of MRMs designed and tested for the detection of 53 proteins in human plasma or serum

Protein name, tryptic peptide sequence, retention time in experiment D, peptide mass (MS1, *m/z*), singly charged fragment MS2 (*m/z*), mean peak area values, and CVs for 10 replicate analyses across five experiments are shown.

044			۵	ARM			Mean Peak Areas					CV (%)					
MRM Count	Protein	Peptide Sequence	RT in D	Best MRM	SIS	MS1/MS2	Α	в	с	D	Е	A	в	с	D	E	
1	Afamin	DADPDTFFAK	21.5	X		563.8 / 825.4						8	9	10	4	- 6	
2						563.8 / 940.4						7	15	17	10	10	
3	Alpha-1-acid glycoprotein 1	NWGLSVYADKPETTK	19.7			570.3 / 1052.5	2.3E+05	5.9E+04	1.9E+05	1.6E+05	3.1E+05	6	12	9	8	17	
4	57 1			х		570.3 / 575.3						3	13	7	8	16	
5					Х	575.6 / 1068.5						11	21	11	25	16	
6	Alpha-1- antichymotryp sin	EIGELYLPK	22.4			531.3 / 633.4	5.0E+05	1.2E+05	7.5E+05	4.3E+05	9.9E+05	5	6	6	3	5	
7				Х		531.3 / 819.5						2	5	7	2	5	
8	Alpha-1B-		27.7		Х	535.3 / 827.5						5	12	3	3	6	
9 10	glycoprotein	LETPDFQLFK	27.7	x		619.4 / 995.5 619.4 / 894.5						4	7	11 7	6	14	
11	Alpha-2-	LGNQEPGGQTALK	12.6	x		656.8 / 771.4						4	20	10	6	8	
	antiplasmin																
12						656.8 / 900.5						9	38	19	15	10	
13 14	alpha-1-	DTEEEDFHVDQVTTVK	17.4		X	660.8 / 779.4 631.3 / 790.4						6 20	20 33	7	7	23	
14	antitrypsin	DIEEDINVDQVIIVK	17.4			051.57750.4	0.52+05	2.92+05	5.52+05	5.02+05	1.02+04	20	55	50	74	25	
15						631.3 / 889.5						11	33	33	29	13	
16	alpha-2- macroglobulin	LLIYAVLPTGDVIGDSAK	36.5	Х		923.0 / 1059.5	4.1E+05	5.3E+04	1.6E+04	4.2E+05	4.3E+05	17	20	19	6	8	
17	macroglobulin					923.0 / 1172.6	1.5E+05	1.9E+04	6.2E+03	1.4E+05	1.4E+05	16	31	20	10	8	
18	Angiotensinog	ALQDQLVLVAAK	23.5			634.9 / 956.6						10	56	8	8	13	
	en						2 05 . 04					-	- 4				
19 20					v	634.9 / 713.5 638.9 / 964.6						7 9	51 17	16 7	10 10	16 8	
21		PKDPTFIPAPIQAK	19.1		^	508.3 / 724.4						14	17	8	15	11	
22				х		508.3 / 556.4						10	19	6	14	15	
23		DDLYVSDAFHK	19.2			437.2 / 803.4	9.1E+04	2.1E+04	1.4E+05	2.6E+04	4.2E+04	5	6	10	13	6	
24	III			х		437.2 / 704.3	3 3E+05	7 5E+04	5 1E+05	9 3E+04	1 5E+05	5	5	7	7	5	
25				~	х	439.9 / 811.4						21	21	15	22	17	
26	Apolipoprotein	ATEHLSTLSEK	12.1	Х		405.9 / 664.4						4	4	7	5	12	
27	A-I					405.9 / 777.5	1.9E+06	4.0E+05	3.8E+06	5.5E+05	1.3E+06	5	4	6	5	12	
28			12.1		Х	408.5 / 672.4						11	8	14	9	9	
29	apolipoprotein A-II precursor	SPELQAEAK	12.1			486.8 / 546.4						19	41	7	5	10	
30 31	Apolinoprotoin	SLAPYAQDTQEK	13.8	X		486.8 / 659.4 675.8 / 982.4						19	24	5 15	4	<u>11</u> 9	
21	A-IV	SLAPTAQUIQEK	12.0	^		0/5.0/902.4	7.4E±04	1.00+04	1./E+04	0.12+04	1.00+05	0	20	15	5	9	
32						675.8 / 1079.5	1.8E+05	3.6E+04	4.3E+04	1.3E+05	3.4E+05	6	26	11	9	6	
33	Apolipoprotein B-100	FPEVDVLTK	22.8			524.3 / 803.5	5.1E+04	6.8E+04	4.2E+05	1.4E+05	3.0E+05	7	7	7	3	5	
34	0 100			х		524.3 / 674.4	4.7E+04	5.8E+04	3.8E+05	1.2E+05	2.7E+05	6	5	6	5	7	
35					х	528.3 / 811.5						9	9	4	2	4	
36		TEVIPPLIENR	22.9			640.8 / 838.4						5 3	11 8	5	6 5	7	
37 38	Apolipoprotein	TPDVSSALDK	14.9	X		640.8 / 741.4 516.8 / 620.3						12	18	11 27	10	10	
50	C-I lipoprotein		1 115	~		51010 / 02015	1102101	,122100	2.52.101	LILLIUI	5.52.101		10		10	Ŭ	
39						516.8 / 719.4						17	23	17	12	9	
40	Apolipoprotein C-II	STAAMSTYTGIFTDQVLS	36.2			745.1 / 1149.7	2.0E+04	4.4E+03	6.8E+03	8.1E+03	1.7E+04	25	41	26	22	17	
	lipoprotein	VLK															
41	hpoprotein					745.1 / 1002.6	1.9E+03	2.5E+03	9.2E+03	5.0E+02	5.3E+03	40	35	30	101	29	
42	Apolipoprotein C-III	DALSSVQESQVAQQAR	17.8			858.9 / 1144.6	1.0E+05	9.5E+04	1.5E+03	1.2E+05	3.6E+05	5	50	8	11	5	
43				х		858.9 / 1417.7	1.7E+04	1.7E+04	2.6E+04	2.2E+04	6.4E+04	16	12	14	11	8	
44	Apolipoprotein	LGPLVEQGR	15.5			484.8 / 701.4	5.6E+04	2.5E+04	8.7E+03	5.7E+04	1.5E+05	10	27	16	8	6	
45	E			Y		484.8 / 588.3	1 6E±05	5 5E±04	2 0E±04	1 3E±05	3 56±05	2	17	8	4	4	
45	Beta-2-	ATVVYQGER	12.4	X		484.8 / 588.3						9	28	8	6	4	
	glycoprotein I											2		5	-		
47				Х		511.8 / 751.4						8	15	7	7	6	
48		EHSSLAFWK	19.3			552.8 / 838.5						10	26	37	16	19	
49 50					х	552.8 / 664.4 556.8 / 846.5						27 16	66 28	56 21	29 33	17 24	
51	C4b-binding	LSLEIEQLELQR	27.0			735.9 / 915.5						12	37	30	22	12	
	protein alpha chain																

TABLE II—continued

÷			۵	MRM			Mean Peak Areas						CV (%)					
MRM Count	Protein	Peptide Sequence	RT in	Best	SIS	MS1/MS2	A	в	с	D	Е	A	в	с	D	E		
52						735.9 / 1028.6	9.7E+03	2.0E+03	1.1E+03	6.5E+03	4.4E+04	16	61	45	26	11		
53	Ceruloplasmin	EYTDASFTNR	14.9			602.3 / 624.3	3.5E+05	7.8E+04	6.4E+04	4.1E+05	1.1E+06	5	19	7	5	4		
54				х		602.3 / 695.3						6	18	5	4	4		
55	Clusterin	LFDSDPITVTVPVEVSR	28.5	V		937.5 / 1296.7						19	13	19	8	8		
56 57	Coagulation	DPPSDLLLLK	26.7	Х		937.5 / 686.4 555.8 / 898.6						19 18	11 26	13 50	6 19	7		
	factor V	DFF3DLLLLK	20.7		V													
58 59	Coagulation	VVGGLVALR	19.7		X	559.8 / 906.6 442.3 / 784.5						14	12	27	4	4		
	factor XIIa heavy chain	VVGGLVALK	19.7										10	27	0	,		
60				X		442.3 / 685.4						6	8	8	4	6		
61	Complement C3	TGLQEVEVK	15.0			501.8 / 731.4	1.7E+06	3.8E+05	5.7E+03	1.6E+06	4.0E+06	3	35	6	6	6		
62				X		501.8 / 603.3						2	35	5	6	6		
63	Constant	TOUR	16.2		Х	505.8 / 739.4	the second s		and the second se	and the second se	A REAL PROPERTY OF THE REAL PR	11	13	10	14	6		
64	Complement C4 gamma chain	IIQVLHFIK	16.2			362.9 / 645.4						9	20	9	5	12		
65				Х	V	362.9 / 744.4						9	18	12	4	7		
66 67	Complement		20.9	Y	Х	365.6 / 653.4 557.8 / 629.4						15	29	9	10	13		
	C4 beta chain	VGDTENENER	20.9	X								4	4		Ĩ	5		
68	Complement		20.2	V		557.8 / 843.5						8	11	7	10	5		
69	Complement C9	AIEDYINEFSVR	28.3	×		728.5 / 1271.6	5.2E+04	9.6E+03	2.8E+04	2.3E+04	1.5E+05	/	11	14	10	/		
70	C					728.5 / 1027.5	3.0E+04	6.1E+03	1.7E+04	1.3E+04	8.5E+04	9	20	24	10	7		
71	Complement factor B	EELLPAQDIK	19.0	Х		578.4 / 671.4						3	6	5	3	3		
72	Tactor D					578.4 / 784.5	2.7E+05	4.5E+04	3.0E+05	2.4E+05	6.1E+05	4	5	11	2	4		
73	Complement factor H	SPDVINGSPISQK	16.3	Х		671.4 / 830.4						8	22	14	10	8		
74						671.4 / 572.3	4.4E+04	8.5E+03	8.4E+03	2.6E+04	7.4E+04	7	23	17	7	10		
75	Fibrinogen alpha chain	TVIGPDGHK	11.7			462.3 / 723.4						22	12	5	4	6		
76					235	462.3 / 610.3						71	9	3	5	8		
77		COLOUTINT	147	v	х	466.2 / 731.4						59	16	5	10	7		
78 79		GSESGIFTNTK	14.7	X		570.8 / 780.4 570.8 / 867.5						4	17 17	6 6	4 5	5		
80	Fibrinogen beta chain	QGFGNVATNTDGK	13.5	Х		654.8 / 706.3						4	16	4	6	5		
81	Deta chain					654.8 / 805.4						4	31	5	12	10		
82 83	Fibrinogen	DTVQIHDITGK	15.4		X	658.8 / 714.3 409.5 / 670.4						4	<u>16</u> 4	4	6	5		
00	gamma chain	DIVQINDIIGK	15.4			409.5 / 6/0.4	2.76+05	7.1E+04	1.00+00	2.20+04	5.4E+04	4	4	0	11	/		
84	guinna chuin			х		409.5 / 533.3	2.6E+05	7.6E+04	1.6E+06	4.5E+04	1.1E+05	3	3	8	4	6		
85					Х	412.2 / 678.4	2.3E+04	5.5E+04	2.5E+04	2.6E+04	6.6E+04	13	14	6	17	7		
86	Fibronectin	DLQFVEVTDVK	24.7	х		647.3 / 789.4						7	13	7	4	4		
87		VTWAPPPSIDLTNFLVR	20.2			647.3 / 690.4						9	8	10	5	4		
88 89		VIWAPPPSIDLINFLVK	38.2			642.7 / 977.5 642.7 / 862.5						41 47	37 48	24 25	14 20	11 18		
90	Gelsolin,	TGAQELLR	14.8	х		444.3 / 786.5						6	14	9	6	7		
91	isoform 1					444 2 / 720 4	1 55 1 05	2 55 1 04	1 55 1 04	1 65105	4 25 1 05	6	22	14	F	6		
92	Haptoglobin	VGYVSGWGR	18.2	Х		444.3 / 729.4 490.8 / 562.3						3	23 9	<u>14</u> 5	5	<u>6</u> 5		
93	beta chain					490.8 / 661.3	2 0E±05	8 1E±05	2 1E±06	4 1E±04	9 4F±04	4	11	5	9	5		
94	Hemopexin	NFPSPVDAAFR	23.6	x		610.8 / 959.6						2	6	4	3	5		
95	nemeperan		2010	~		610.8 / 775.3						3	9	5	3	6		
96	Heparin	TLEAQLTPR	16.4	Х		514.8 / 814.4						4	22	8	5	3		
97	cofactor II					514.8 / 685.4	2.9E+05	7.0E+04	1.8E+04	2.5E+05	6.8E+05	5	25	8	5	4		
98	Histidine-rich	DSPVLIDFFEDTER	37.7	Х		841.9 / 1171.5						34	23	22	9	11		
99	glycoprotein					841.9 / 1058.4						33	26	22	10	11		
	Inter-alpha- trypsin inhibitor heavy chain	AAISGENAGLVR	14.9	Х		579.4 / 902.5						4	17	4	3	3		
101						579.4 / 629.4						4		6	7	4		
102	Inter-alpha- trypsin	AFIQLWAFDAVK	38.0			704.9 / 836.4	3.0E+04	5.3E+03	2.8E+03	2.9E+03	2.6E+03	33	51	32	30	39		
103	inhibitor light					704.9 / 949.5	1.5E+04	2.9E+03	7.3E+02	1.2E+03	1.1E+03	34	47	32	56	39		

				RM				Mea	n Peak A	reas			C	(%)	)	
104 105	<b>Protein</b> Kininogen	<b>Peptide Sequence</b> TVGSDTFYSFK	<b>D</b> <b>B</b> 21.5	imes Best MRM	SIS	<b>MS1/MS2</b> 626.3 / 1051.4 626.3 / 994.5						A 6 9	<b>B</b> 6 10	c 6 18	<b>D</b> 5 8	<b>E</b> 5 35
105 106 107 108	L-selectin	AEIEYLEK	16.9	x	×	497.8 / 794.4 497.8 / 681.3 501.8 / 802.4	1.9E+04 1.2E+04	6.0E+03 3.4E+03	5.6E+03 2.5E+04	1.7E+04 1.3E+04	4.7E+04 3.3E+04	16 17 6	75 22 54	19 25 6	14 22 8	6 11 6
109	Plasma retinol-binding protein precursor	YWGVASFLQK 9	35.1	Х	~	599.8 / 693.4	9.5E+04	1.1E+04	1.3E+04	4.8E+04	6.1E+04	4	17 20	25	17	18
111 112 113	Plasminogen	LSSPAVITDK	15.3		x	515.8 / 743.4 515.8 / 830.5 519.8 / 751.4	1.8E+05 1.2E+05	8.8E+04 5.4E+04	4.7E+04 3.2E+04	1.7E+05 1.1E+05	4.3E+05 2.7E+05	5 5 7	18 13 25	11 11 8	5 6 5	5
114 115	Duatharanahia	LFLEPTR	19.1	х		438.3 / 615.4 438.3 / 502.3	5.3E+05 2.7E+05	2.2E+05 1.1E+05	1.2E+06 6.4E+05	5.0E+05 2.9E+05	1.2E+06 6.1E+05	7 5	28 4	6 8	3 6	4
116 117 118	Prothrombin	ETAASLLQAGYK	20.2	х	х	626.3 / 879.5 626.3 / 679.4 630.3 / 887.5	3.9E+05 4.4E+04	6.5E+04 1.5E+05	2.3E+05 4.4E+04	2.0E+05 7.7E+04	5.2E+05 2.1E+05	3 4 13	10 9 9	10 7 6	5 6 6	4 4 4
119 120	Serum albumin	LVNEVTEFAK	19.3	х		575.4 / 937.4 575.4 / 694.4						14 14	19 5	4 5	18 8	13 18
121	Serum amyloid P- component	VGEYSLYIGR	21.3			578.8 / 1057.5	2.1E+04	1.6E+04	2.7E+04	1.0E+05	5.5E+04	7	9	17	7	11
122	110			Х		578.8 / 871.5						11	6	12	7	10
123 124	Transferrin	EDPQTFYYAVAVVK	20.3	Х		815.4 / 1160.6 815.4 / 1288.7	8.9E+02	3.0E+04	2.0E+04	4.4E+02	6.5E+02	30 66	11 10	5 10	31 59	24 98
125 126	Transthyretin	AADDTWEPFASGK	22.3	х		697.8 / 921.4 697.8 / 606.4						5 6	8 3	16 14	8 3	6 7
127	Vitamin D- binding protein	THLPEVFLSK	19.7	Х		585.8 / 819.5						9	8	27	12	17
128 129	Vitamin K- dependent protein C	WELDLDIK	20.1			585.8 / 932.5 516.3 / 716.4						13 61	14 31	34 25	12 20	<u>14</u> 17
130 131	protein e				х	516.3 / 603.3 520.3 / 724.4						98 9	41 12	30 32	17 24	23 32
132 133	Vitronectin	DVWGIEGPIDAAFTR	36.4			823.9 / 947.5 823.9 / 890.5	6.7E+04 3.5E+04	8.1E+04 4.1E+04	4.7E+04 2.6E+04	1.4E+05 7.7E+04	2.7E+05 1.4E+05	36 35	27 29	16 22	7 6	9 8
134 135		FEDGVLDPDYPR	22.2	Х		711.9 / 875.4 711.9 / 1031.5						4 5	6 9	7 14	4 9	6 9
136	Zinc-alpha-2- glycoprotein	EIPAWVPFDPAAQITK	36.2	Х		891.9 / 1087.7						40	29	12	13	15
137						891.9 / 728.4	3.8E+03	1.7E+04	9.8E+04	7.3E+03	1.4E+04	43	28	17	22	19

TABLE II—continued

spin column). In comparing average peak areas obtained in analyses of digests of whole and depleted samples, we found substantial reductions in albumin (1.3E+08 reduced to ~1E+04), transferrin (1.5E+05 reduced to ~5E+03), and haptoglobin (4.6E+06 reduced to ~1E+05).  $\alpha_1$ -Antitrypsin was not detected reliably in any of the runs (presumably a bad peptide choice for MRM assay), and so its removal could not be confirmed. We did not incorporate MRMs to measure the two immunoglobulins subtracted by the MARS column.

Multiple measurements of an MRM would be expected to improve CVs, and we thus examined whether the sum of the two fragment MRMs measured separately for 59 of the peptides exhibited a smaller CV across 10 replicate runs than the individual MRMs. As shown in Table III, the average CV for the summed MRMs across 59 peptides is 1–3% lower than the averages of either individual MRM. If the summed CV is

compared with the lower of the two fragment CVs for each MRM, the average reduction in CV in experiments A–E ranges from +0.7% to -0.1%. These small improvements in CV come at the cost of doubling the measurement time (or halving the number of peptides monitored) and thus are unlikely to be useful in routine operation for improving reproducibility.

The relationship between CV and peak area, shown in Fig. 3 for experiments A–E, indicates that peak areas below 1E+04 are unlikely to yield acceptable CVs (*i.e.* below 10%). A cutoff of 1E+04 corresponds to a signal-to-noise ratio of ~10, which is consistent with the quantitative requirement of a S/N ratio of 10 for a reported lower limit of quantitation. The highest peak areas measured (albumin peptides in whole plasma digest samples) are above 1E+08, demonstrating a maximal working dynamic range of >4 orders of magnitude above this cutoff.

Calibration with Internal Standard Peptides - A set of 13 SIS peptides having the same sequences as 13 of the MRMmeasured tryptic peptides were used to assess the reproducibility of the quantitative measurements (five of the 18 SISs measured by MRM yielded unusable quantitative information because the peak area signal of either the standard or its cognate sample digest peptide fell below 1E+04). The SIS peptides were labeled with [U-13C6,U-15N2]lysine during synthesis from a synthetic DNA template in an in vitro transcription/translation system and were cleaved from the polySIS protein by trypsin. Samples spiked with SIS peptides were analyzed in experiments D and E (Table I) at relative peptide loadings of 1× and 3.3×, respectively (SIS peptides at relative loadings of  $1 \times$  and  $3 \times$ ). Measured peak areas of MRM peptides showed an average 2.7× difference between these experiments. Ratios between the natural sample-derived (Nat) and labeled (SIS) versions of these peptides were extremely similar in the two experiments (Fig. 4) and maintained a linear relationship ( $R^2 = 0.998$ ) over almost 3 orders of magnitude in concentration. Because the Nat:SIS ratios combine the variance of two measurements, the CVs of the ratios are typically

TABLE III Improvement of CV obtained by combining peak areas of two fragment MRMs versus keeping them separate

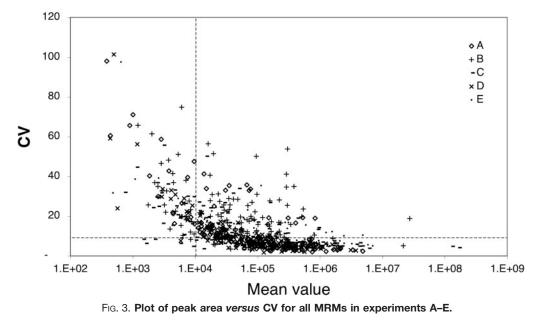
Avg, average	e; frag(s),	fragment(s).	
--------------	-------------	--------------	--

Exporimont	Avg CV								
Experiment	Sum of frags	Frag 1	Frag 2						
		%							
А	10.5	11.8	14.8						
В	16.2	20.0	19.4						
С	11.0	13.0	14.4						
D	8.0	9.4	12.3						
E	8.5	9.4	11.9						

slightly higher than CVs of the component measurements (80 of 130 comparisons).

Sensitivity-Two proteins with relatively low normal concentrations in plasma were clearly detected among the MRMs tested: L-selectin and fibronectin. The soluble form of Lselectin is a 33-kDa protein present in plasma at a normal concentration of  $\sim$ 0.67  $\mu$ g/ml (29) or 20.3 pmol/ml. Fibronectin is a 260-kDa protein present in plasma at a normal concentration of  $\sim$ 300  $\mu$ g/ml (30) or 1200 pmol/ml. Given that an amount of digest corresponding to 0.01  $\mu$ l of plasma was loaded on column in experiment D, these peptides would be expected to be present on column at ~200 and 12,000 amol, respectively. In the case of L-selectin we had spiked a SIS peptide at 2.0 fmol and thus could determine that the natural (sample-derived) peptide was present at 0.1 times the amount of SIS (single point quantitation), yielding a measured 200 amol and implied plasma concentration of 0.6–0.67  $\mu$ g/ml in good agreement with expectation. CVs for fibronectin in experiments D and E were 4 and 4%, respectively, and for L-selectin were 22 and 11%, respectively, presumably reflecting the fact that L-selectin was near the lower limit ( $\sim$ 1E+04) for high quality detection.

Six of the 53 selected target proteins were not reliably observed, however. Three are probably explained by low normal abundances: we did not obtain a reproducible signal for the selected peptides from coagulation factor V, vitamin K-dependent protein C, or C4b-binding protein. There were also instances in which peptides from more abundant proteins were not reliably detected: the inter- $\alpha$  trypsin inhibitor light chain (despite the fact that a peptide from the heavy chain of this protein gave a good quality MRM), apolipoprotein C-II, and  $\alpha_1$ -antitrypsin. In these latter cases, the problem is most likely poor choice of peptides: numerous alternative peptides exist for both the inter- $\alpha$  trypsin inhibitor light chain and



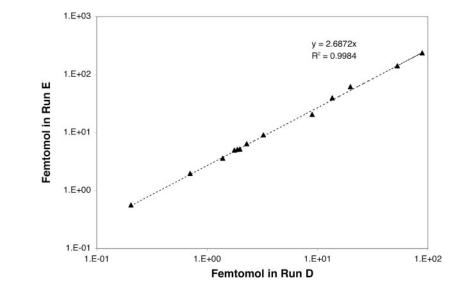


Fig. 4. Comparison of computed amounts (based on ratios between Nat and SIS peptides) for 13 peptides in experiments D and E. Data are included for those ratios where both numerator and denominator peak areas were >1E+04 in experiment D (13 peptides).

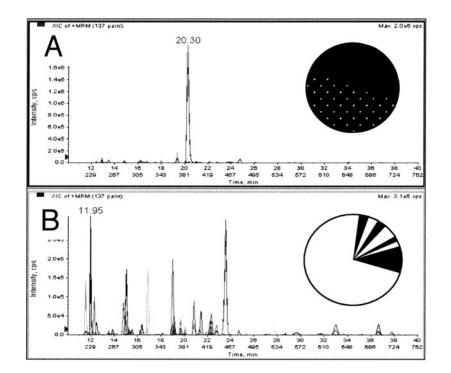
 $\alpha_1$ -antitrypsin, but for small proteins, such as apolipoprotein C-II, there may be no better alternative, and additional enrichment of these peptides will be necessary.

In general, depletion of the most abundant proteins using the Agilent MARS column improved the performance of MRMs for non-subtracted proteins substantially. This effect appeared to be due to improved detection sensitivity and improved chromatographic peak shape (achieved by decreasing the total peptide loaded by ~6-fold), both of which contribute to improved MRM peak signal-to-noise and lower CVs. Fig. 5 illustrates the benefit of depletion in removing the albumin peptide (major peak in Fig. 5*A*) and thus boosting the minor peaks in the depleted sample (Fig. 5*B*). At very high loading of undepleted plasma digest, we noticed large shifts in peak retention times, but at loadings in the region of our nominal load the effect of high abundance peptides on MRM retention times was minor.

Reproducibility of Immunodepletion and Tryptic Digestion-In a sixth experiment (F series), we performed MARS depletion on two aliquots of the same plasma sample and then separately digested two aliquots of each depleted sample (total of four samples; e.g. F1\_2 refers to the second digest of the first depletion). Four replicate runs of the 137 MRMs were carried out for each sample in randomized order to avoid any sequence effect. Fig. 6 compares the mean peak areas of two digests of a single depleted sample (Fig. 6A, F1\_1 versus F1\_2) and two depletions (Fig. 6B, F1\_1 versus F2\_1). The peak area values span 4 orders of magnitude of which  $\sim$ 2.5 orders is usable (*i.e.* beginning at peak area >1E+04, the approximate cutoff below which CVs become large). Duplicate digests show excellent comparability ( $R^2 = 0.995$  and 0.998 for F1\_1 versus F1\_2 and F2\_1 versus F2\_2, respectively). Duplicate depletions (which necessarily include the effects of different digests as well) are only slightly worse (e.g.  $R^2 = 0.989$  and 0.991 for F2\_1 versus F1\_1 or F2\_2 versus F1\_2, respectively).

The Density of Signals in MRM Space: Behavior of Random MRMs-Most of the MRMs we designed appeared to detect only a single peak during the LC run of a complex digest such as depleted plasma (e.g. Fig. 1): whereas 73% had a robust peak (defined as area >32,000) corresponding to the target peptide analyte, only about 8% had a second peak meeting the same peak area criterion (histograms in Fig. 7). We therefore attempted to confirm that the density of peptide peaks in "MRM space" was indeed low (equivalent to high MS/MS detector specificity relative to sample complexity) by examining two types of randomized MRMs in the same depleted plasma digest sample. In a first set, 100 MRMs were generated with "parent" masses randomly distributed over the mass range of real peptides used in the 137 designed MRMs and "fragment" masses randomly distributed between the parent mass and the maximum fragment mass among the designed MRMs ("random MRMs"). Of the 100 random MRMs, only six showed a peak with area >32,000, and none of these MRM peaks produced MS/MS spectra that led to a protein identification when searched with Mascot against Swiss-Prot. A second set of 131 random MRMs was generated by randomly pairing parent and fragment masses from the set of designed MRMs detectable in plasma, excluding those cases where the fragment mass was lower than the parent ("random combination MRMs"). By using real peptide and fragment masses, these MRMs avoided any potential bias arising from the tendency of real peptide masses to cluster around integral masses (the mass defect). In this second set, about 12% of the MRMs exhibited a peak with peak area >32,000, and once again none of these peaks gave MS/MS spectra yielding a protein identification. All the peaks observed in the random MRM sets occurred late in the LC gradient (after 100 min) after the elution of a large majority of the designed plasma protein MRMs. These results suggest that the density of quantifiable features in MRM space at our current sensitivity, even for a very complex peptide sample

FIG. 5. Total ion current profiles across the chromatographic peptide separation for digests of undepleted (A, whole) and depleted (B) plasma from experiments C and D, respectively. Pie charts represent the protein composition of the samples: whole plasma contains >50% albumin (stippled region), whereas the proteins remaining after MARS depletion include fibrinogen,  $\alpha_2$ -macroglobulin, complement C3, and all other lower abundance proteins (four remaining pie slices ordered clockwise from 12 o'clock), showing the proportion of protein removed by depletion (white segments in B). The major peptide in A at 20.3 min is derived from albumin (whose abundance is shown by the *black* pie segment with white dots). cps, counts per second.



and using unit resolution in both mass analyzers, is only 6–12% of which a minority may be canonical tryptic peptides. The distribution of peak areas observed for random MRMs closely matches the distribution for second (non-target peptide) peaks in our designed MRMs, indicating that these additional peaks represent a random background.

Selection of Best MRMs for Plasma Proteins-From the 119 tested MRMs for sample peptides, we selected the best fragment ions (from 59 cases of two tested fragment ions) and the best peptide (from seven cases where two peptides were tested per protein). A usable MRM resulted for 47 of the 53 initial proteins (indicated by an X in the Best MRM column in Table II). Each of the 47 peptide sequences was verified as unique in the human proteome (represented by the Ensembl peptides) and occurred only once in the target protein. Three of the peptides (representing antithrombin III, apolipoprotein E, and vitamin K-dependent protein C) occur in the mouse homologs as well, and seven (apolipoprotein E, vitamin K-dependent protein C, complement C4  $\beta$  and  $\gamma$ , fibronectin, haptoglobin  $\beta$ , and inter- $\alpha$  trypsin inhibitor heavy chain) occur in rat homologs (all the other human sequences did not occur in the other species' Ensembl peptides).

Of the final 47 MRMs, 12 were contributed by the *in silico* approach leading to the 30 polySIS peptides, and one (hemopexin) was contributed by an earlier *in silico* effort (19): these are indicated in Table II by an X in the SIS column for the associated Lys-labeled internal standard. A total of eight *in silico* selections were replaced by better performing peptides from the same target protein as a result of experimental testing (four before and four after selection of the 137 MRMs), two subsequently failed and have not yet been replaced, and

eight were dropped before testing because of expected insufficient abundance. Thus 13 *in silico* selections survived, whereas 10 were replaced in testing.

Of the six failures (see above), we expect three proteins to be within the concentration range that should be measurable and are selecting substitute peptides for these. The distribution of CVs for the 47 best MRMs is shown in Fig. 2*B*: in experiment D, 40 of these had CVs below 10%, and 19 had CVs below 5%.

### DISCUSSION

We designed and characterized mass spectrometric MRM assays for tryptic peptides representing 53 high-to-medium abundance proteins of human plasma. The resulting panel of 47 successful MRMs, having within-run CVs ranging from 2 to 22% in depleted plasma, should be useful in the exploration of protein-disease relationships and expression control in clinical serum and plasma specimens. The fact that MRM assays, once developed, can be implemented on triple quadrupole (QqQ)-MS instruments, widely used for precise quantitation of small molecules at high throughput, offers the prospect of data collection across very large clinical sample sets. Our approach exemplifies the development of targeted specific assays needed to enable the statistical validation of proposed biomarkers in plasma ("candidate-based proteomics" (4)).

Our initial attempts to select usable tryptic peptides by purely *in silico* means were reasonably successful as approximately half of the peptides chosen (13 of 23) produced acceptable MS signals in plasma. Although the prediction of ionization properties of tryptic peptides can be expected to

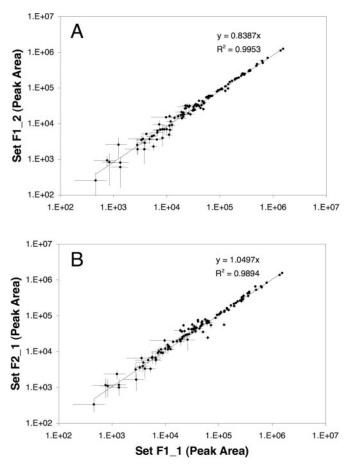


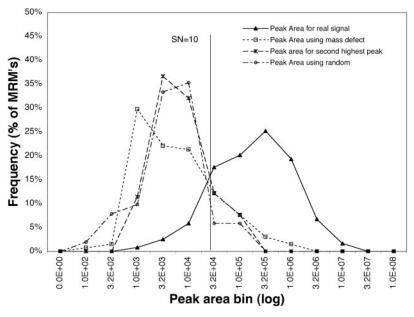
Fig. 6. Comparison of mean peak areas values for all MRMs in two replicate digests (A) and two depletion runs (and subsequent digests) (B). *Error bars* show 1 standard deviation computed from replicate runs.

improve substantially in the future, we used experimental MS/MS data, in combination with computational methods, to select more successful target peptides. Two experimental methods proved particularly useful. High abundance peptides were detected in conventional LC-MS/MS data-dependent full scan MS experiments in which a subset of high signal peptides seen in MS1 are subjected to MS/MS. Lower abundance peptides were detected by constructing lists of candidate MRMs to all appropriately sized predicted tryptic peptides from a target protein and then characterizing any detected MRM peaks by MS/MS (the MIDAS workflow in which MRM methods are designed using a specifically designed script within the Analyst software). Because MRMs are typically more sensitive than full scan survey MS for detection of very low abundance components, the MIDAS approach allowed us to find successful MRMs for more lower abundance peptides and represents the preferred strategy for MRM design going forward. This process was facilitated by the combination of high sensitivity triple guadrupole MRM and ion trap MS/MS scan capabilities on the hybrid triple quadrupole linear ion trap 4000 Q TRAP mass spectrometer. A general process for MRM design using these sources of information is shown in Fig. 8 in which *bold arrows* indicate the preferred approach. The final step in the process, selection among alternative candidate MRMs for a given protein based on measured CVs in replicate runs, is recommended because of the significant differences in CV observed among MRMs with similar signal strength: CV appears to be a characteristic of a peptide to some extent separate from ion current.

Despite the complexity of plasma digests (particularly those of depleted plasma where a small number of superabundant peptides have been removed), most MRMs exhibited only a single peak across the peptide LC chromatogram. This observation is consistent with the low density of peaks in two sets of randomly distributed MRMs measured in depleted plasma digests and demonstrates the exquisite specificity of the two-stage QqQ-MS selection process used as the detector. Nevertheless the existence of secondary peaks (whether or not they are actually tryptic peptides) in a small subset  $(\sim 10\%)$  of MRMs indicates that chromatographic elution time is an important additional factor in providing the absolute analyte specificity desired in these assays. Following more extensive experience with specific MRMs and any variation that may occur in them due to shifts in peak elution times, it may be possible to establish a set of plasma MRMs that are truly free of secondary peaks and that could therefore potentially be measured using short LC separations (constrained only by increases in ion suppression effects as peptides crowd together).

The unambiguous measurement of a peptide derived from L-selectin (normal concentration of ~0.67) suggests that proteins present in normal plasma at  $\sim 1 \ \mu$ g/ml are measurable by MRM in depleted and undepleted samples with minimal upfront sample fractionation. Different tryptic peptides from the same protein can produce ion currents differing by factors of at least 1E+03 in LC-MS/MS experiments, excluding the peptides that are not detected at all. This variation is due to multiple factors, including propensity to ionize in the electrospray source, coincidence in elution time with other easily ionizing peptides, efficiency of release during tryptic digestion, and presence of unrecognized post-translational modifications arising due to biology or during sample preparation. Although we do not yet know the frequency of high efficiency peptides among the tryptic products of each protein, we believe that numerous additional plasma proteins can be measured down to  $\sim 1 \ \mu g/ml$  using MRMs and estimate that 50-100 proteins may be added to our list as we extend this effort to additional candidates from the plasma component database (2, 20).

Given that albumin, the most abundant protein in plasma, is measurable by MRM in undepleted plasma in the same experiment that detects L-selectin, the dynamic range of quantitatively measurable proteins appears to be  $\sim 1 \ \mu$ g/ml to 55 mg/ml or  $\sim 5E+04$ . This dynamic range of 4–5 orders of magnitude is consistent with typical quantitation experiments



it is easily possible to select lower signal peptides from high abundance proteins (and high signal peptides from low abundance proteins) to diminish the dynamic range requirement at the peptide level. As long as stable isotope-labeled internal standards are used, the quantitative information contained in Nat:SIS ratios should be unaffected by peptide choice.

The reproducibility of peptide MRM measurements was impressive in many cases. For the three experiments analyzing depleted plasma digests, an average of 78% of the 47 best MRMs had within-run CVs of 10% or less, and 36% had CVs less than 5%. A number of MRMs gave low CVs across five experiments (i.e. in depleted and whole plasma digests at widely varying total loads) including peptides from hemopexin (CVs ranging between 2 and 6%), vitronectin (4-7%), kininogen (5-6%), complement factor B (3-6%), complement C4 (3-5%), apolipoprotein B100 (5-7%), antithrombin III (5-7%), and  $\alpha_1$ -antichymotrypsin (2–7%). These precision values are comparable to many small molecule MRM measurements and approach the results of conventional immunoassays used in clinical diagnostics (32). They demonstrate clearly that peptide MRM measurements can be used in high precision quantitative assays. Based on reproducibility at this level, it appears likely that useful comparative data across samples might be obtained by comparing simple peak area measurements (i.e. without SIS standardization) and that relatively small quantitative differences could be detected (differences of 20% would be >3 standard deviations from the mean for many of the best MRMs and thus very easily detected). The observation that individual peptides showed consistently higher or lower CVs suggests that quantitative precision depends on characteristics of the peptide and perhaps its environment of co-eluting peptides. Therefore it will be useful in future work to select MRM peptides based on CV across replicate runs in addition to the characteristics normally con-

Fig. 7. Histograms of peak areas obtained with four sets of MRMs in a digest of depleted plasma.

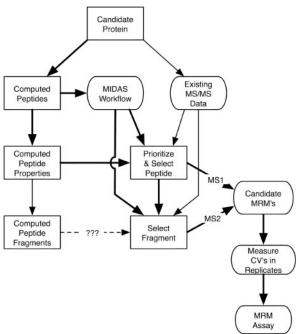


FIG. 8. Schematic diagram of a process for selecting MRMs. *Thicker arrows* represent the preferred process. *Square-cornered boxes* represent informatics components, and *rounded-corner boxes* are experimental steps.

performed using MRM on this instrument and covers almost half the logwise dynamic range of currently known and measured proteins in human plasma (1). The dynamic range of peak area measurements appears to be similar, extending from  $1.3E+08 \pm 5\%$  for undepleted albumin to less than 1E+04(the approximate cutoff below which CVs become >10%). In this context it is important to note that measurement of proteins over this dynamic range does not require measuring peptides at such a broad range of signal intensities because sidered (e.g. ion current and chromatographic peak shape).

MRMs designed for absolute quantitation using a set of 13 SIS peptides (with sequences identical to 13 of the selected tryptic peptides) performed well. When spiked at known loading these provided internal standards for quantitation and resulted in a measured value for L-selectin in almost exact agreement with the literature value for normal subjects (0.67  $\mu$ g/ml). This result is encouraging but not definitive as we did not measure target protein concentrations in our sample by independent methods, and the individual SIS peptides (derived by digestion of a quantitated precursor protein) were not individually quantitated. Thus not all of our endogenous: standard ratio measurements agreed so well with literature values. In addition, single point concentration curves as used here do not provide the same accuracy in quantitation as multiple point curves. Nevertheless the extreme reproducibility ( $R^2 = 0.991$ ) of the measured ratios (and implied absolute peptide concentrations) between experiments D and E (Fig. 4) across almost 3 orders of magnitude in peak area and 3-fold difference in total peptide load are striking. To effectively standardize peptides (and proteins) over wider ranges, a more sophisticated strategy will be utilized using smaller equimolar groups of peptides selected to represent each decade of peptide concentration (i.e. one polySIS product or a set of equimolar SIS peptides for each decade of the abundance scale). In addition, the completeness of polySIS digestion, and thus the relative stoichiometry of the resulting SIS peptides, must be rigorously characterized.

Depletion of the highest abundance plasma proteins using an immobilized antibody column proved to be very useful in measuring our MRMs. We expected to be able to load a digest of ~6 times as much depleted plasma as undepleted plasma because these two samples would contain approximately equal amounts of total peptides, and this was confirmed. We also confirmed that the reference loading chosen (~110 ng of peptides derived from digestion of 10 nl of depleted plasma) was near but not above the limit for good quality LC peak shape and reproducible chromatography in a nanoflow system. However, we had few if any MRMs that were detectable in the depleted sample but not the undepleted sample. The major difference emerged instead in the lower CVs of peaks measured in the depleted sample due presumably to the decreased level of competing peptides and consequent higher signal to noise. In future extensions of this method to lower abundance proteins, it is likely that depletion will expose numerous targets not otherwise detectable. Depletion significantly reduced the levels of albumin, transferrin, and haptoglobin as expected.

Highly reproducible depletion and tryptic digestion are a necessity as steps in routine sample preparation for MRM analysis. In a small study (experiment F) designed to look at these effects,  $R^2$  values  $\geq 0.995$  were obtained for peak areas in duplicate digests, and  $R^2$  values  $\geq 0.989$  were obtained for duplicate depletion and digestion in the same experimental

set. It thus appears that both depletion and digestion can be carried out with sufficient reproducibility to provide useful measurements and that MRM assays provide an ideal method for assessing sample preparation reproducibility.

The multiplexing capability of LC-QqQ-MS platforms for measuring peptides in complex digests is substantial, providing an opportunity to measure large panels of proteins accurately in each run. Based on the performance of the present set of 137 MRMs, which were all monitored continuously across the entire LC gradient as 18-ms sequential measurements, it is clear that 100-200 MRMs might be used routinely to measure peptides in long LC gradients. However, given reproducible chromatographic elution times, it is possible with existing systems to measure each MRM only during a short time window when the peak is expected to occur (e.g. a window of 10% of total run length given an average 2-2.5% CV in peak elution time measured in our experiments D and E). Once this approach is implemented, based on extensive knowledge of elution time and column reproducibility, and provided that MRMs are selected that do not cluster too much in elution time, substantially more MRMs could be used in a single LC MRM experiment.

An additional important consideration for throughput of MRM measurements is the duration of the chromatography run. In our replicate experiments D and E, a 30-min gradient was used that led to a total cycle time (including intersample wash) of 75 min. It would clearly be advantageous if this could be reduced, allowing more samples to be run per day. The extreme analyte specificity indicated by the low density of peaks in MRM space suggests that most of our MRMs should perform well with less benefit from chromatographic separation, and the ability to focus the MRM measurements in discrete time windows may allow more MRMs to be brought closer together in elution time without sacrificing the required multiple measurements across each peak. Hence we expect that substantial improvements in run time should be possible, probably in conjunction with a shift to higher flow rate (e.g. capillary flow) systems providing increased robustness in routine operation. Higher flow rate LC systems have significantly lower ( $\sim 10 \times$ ) absolute sensitivity than nanoflow, but in the current application it would be easily possible to load  $10\times$ more sample (*i.e.* 100 nl) as sample at these loadings is rarely limiting.

Considered in a broad context, MRM assays appear to have several advantages in addition to those described above. 1) The instrumentation used to measure peptide MRMs is very similar to that used in existing robust platforms for high throughput quantitative measurement of drug metabolites in plasma, for the detection of inborn errors of metabolism in newborns and for pesticide analysis. A large base of compatible instrumentation and expertise thus already exists. 2) Because the target proteins are detected via tryptic peptide surrogates after sample denaturation and digestion, peptide MRM results should be insensitive to alterations in protein folding and intersubunit associations (factors that can negatively affect immunoassays). 3) MRM results are obtained by integrating discrete, anticipated peaks in simple ion chromatograms using well established and widely used commercial software, and thus the method is not dependent on the much more complex data processing infrastructure required to handle the peak matching and spectral identification challenges of a discovery proteomic pipeline. 4) It is probable, although not yet demonstrated, that the cost per high quality measurement will be lower than for immunoassays (both to create the assay to begin with and to apply it to large sample sets).

Important limitations of peptide MRMs must also be recognized. 1) Although the method is applicable to post-translationally modified peptides, the post-translational modification must be known or specifically hypothesized, and its properties must be designed into the assay up front (25) (and normally cannot be discovered de novo by our approach). 2) Some proteins may not readily produce usable peptides. The protein may be too short and hence produce few candidates to begin with or may, like many immunoglobulin sequences, be too variable. These cases will require further exploration (alternative proteolytic enzymes or sample derivatization procedures). 3) Genetic variants altering a single amino acid in the selected peptide will prevent its determination by the wild-type MRM; hence these must be recognized in advance and designed specifically. Provided these issues are taken into account, none appears to prevent productive use of peptide MRMs for measuring abundances of plasma proteins.

Based on the results presented here, numerous clinically important plasma proteins can be measured by peptide MRMs with precision comparable to current clinical immunoassays. The panel of assays presented should have general use as the nucleus of candidate-based biomarker validation approach and should lead ultimately toward a comprehensive assay platform for all human proteins in plasma.

Acknowledgment—We thank Dr. Tina Settineri (Applied Biosystems) for advice and support.

\* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ To whom correspondence should be addressed: The Plasma Proteome Inst., P. O. Box 53450, Washington, D. C. 20009-3450. Tel.: 301-728-1451; Fax: 202-234-9175; E-mail: leighanderson@ plasmaproteome.org.

#### REFERENCES

- Anderson, N. L., and Anderson, N. G. (2002) The human plasma proteome: history, character, and diagnostic prospects. *Mol. Cell. Proteomics* 1, 845–867
- Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., Tirumalai, R. S., Conrads, T. P., Veenstra, T. D., Adkins, J. N., Pounds, J. G., Fagan, R., and Lobley, A. (2004) The human plasma proteome: a non-redundant list developed by combination of four separate sources. *Mol. Cell. Proteomics* 3, 311–326
- 3. Kuhn, E., Wu, J., Karl, J., Liao, H., Zolg, W., and Guild, B. (2004) Quantifi-

cation of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and  $^{13}\mathrm{C-labeled}$  peptide standards. *Proteomics* **4**, 1175–1186

- Anderson, N. L. (2005) The roles of multiple proteomics platforms in a pipeline for new diagnostics. *Mol. Cell. Proteomics* 4, 1441–1444
- Joos, T. O., Stoll, D., and Templin, M. F. (2002) Miniaturised multiplexed immunoassays. *Curr. Opin. Chem. Biol.* 6, 76–80
- Haab, B. B. (2005) Antibody arrays in cancer research. Mol. Cell. Proteomics 4, 377–383
- Kostiainen, R., Kotiaho, T., Kuuranne, T., and Auriola, S. (2003) Liquid chromatography/atmospheric pressure ionization-mass spectrometry in drug metabolism studies. *J. Mass. Spectrom.* 38, 357–372
- Tai, S. S., Bunk, D. M., White, E. T., and Welch, M. J. (2004) Development and evaluation of a reference measurement procedure for the determination of total 3,3',5-triiodothyronine in human serum using isotopedilution liquid chromatography-tandem mass spectrometry. *Anal. Chem.* 76, 5092–5096
- Ahmed, N., and Thornalley, P. J. (2003) Quantitative screening of protein biomarkers of early glycation, advanced glycation, oxidation and nitrosation in cellular and extracellular proteins by tandem mass spectrometry multiple reaction monitoring. *Biochem. Soc. Trans.* **31**, 1417–1422
- Sannino, A., Bolzoni, L., and Bandini, M. (2004) Application of liquid chromatography with electrospray tandem mass spectrometry to the determination of a new generation of pesticides in processed fruits and vegetables. J. Chromatogr. A 1036, 161–169
- Barr, D. B., Barr, J. R., Maggio, V. L., Whitehead, R. D., Jr., Sadowski, M. A., Whyatt, R. M., and Needham, L. L. (2002) A multi-analyte method for the quantification of contemporary pesticides in human serum and plasma using high-resolution mass spectrometry. *J. Chromatogr. B. Anal. Technol. Biomed. Life Sci.* **778**, 99–111
- Roschinger, W., Olgemoller, B., Fingerhut, R., Liebl, B., and Roscher, A. A. (2003) Advances in analytical mass spectrometry to improve screening for inherited metabolic diseases. *Eur. J. Pediatr.* 162, Suppl. 1, S67–S76
- Streit, F., Armstrong, V. W., and Oellerich, M. (2002) Rapid liquid chromatography-tandem mass spectrometry routine method for simultaneous determination of sirolimus, everolimus, tacrolimus, and cyclosporin a in whole blood. *Clin. Chem.* 48, 955–958
- Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6940–6945
- Barr, J. R., Maggio, V. L., Patterson, D. G., Jr., Cooper, G. R., Henderson, L. O., Turner, W. E., Smith, S. J., Hannon, W. H., Needham, L. L., and Sampson, E. J. (1996) Isotope dilution-mass spectrometric quantification of specific proteins: model application with apolipoprotein A-I. *Clin. Chem.* 42, 1676–1682
- Wu, S. L., Amato, H., Biringer, R., Choudhary, G., Shieh, P., and Hancock, W. S. (2002) Targeted proteomics of low-level proteins in human plasma by LC/MSn: using human growth hormone as a model system. *J. Proteome Res.* 1, 459–465
- Barnidge, D. R., Goodmanson, M. K., Klee, G. G., and Muddiman, D. C. (2004) Absolute quantification of the model biomarker prostate-specific antigen in serum by LC-MS/MS using protein cleavage and isotope dilution mass spectrometry. *J. Proteome Res.* 3, 644–652
- Liao, H., Wu, J., Kuhn, E., Chin, W., Chang, B., Jones, M. D., O'Neil, S., Clauser, K. R., Karl, J., Hasler, F., Roubenoff, R., Zolg, W., and Guild, B. C. (2004) Use of mass spectrometry to identify protein biomarkers of disease severity in the synovial fluid and serum of patients with rheumatoid arthritis. *Arthritis Rheum.* **50**, 3792–3803
- Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., and Pearson, T. W. (2004) Mass spectrometric quantitation of peptides and proteins using stable isotope standards and capture by anti-peptide antibodies (SISCAPA). *J. Proteome Res.* **3**, 235–244
- Anderson, N. L. (2005) Candidate-based proteomics in the search for biomarkers of cardiovascular disease. J. Physiol. 563, 23–60
- Hopp, T. P., and Woods, K. R. (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc. Natl. Acad. Sci. U. S. A.* 78, 3824–3828
- Krokhin, O. V., Craig, R., Spicer, V., Ens, W., Standing, K. G., Beavis, R. C., and Wilkins, J. A. (2004) An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Mol. Cell.*

Proteomics 3, 908-919

- Adkins, J. N., Varnum, S. M., Auberry, K. J., Moore, R. J., Angell, N. H., Smith, R. D., Springer, D. L., and Pounds, J. G. (2002) Toward a human blood serum proteome: analysis by multidimensional separation coupled with mass spectrometry. *Mol. Cell. Proteomics* 1, 947–955
- Craig, R., Cortens, J. P., and Beavis, R. C. (2004) Open source system for analyzing, validating, and storing protein identification data. *J. Proteome Res.* 3, 1234–1242
- Unwin, R. D., Griffiths, J. R., Leverentz, M. K., Grallert, A., Hagan, I. M., and Whetton, A. D. (2005) Multiple reaction monitoring to identify sites of protein phosphorylation with high sensitivity. *Mol. Cell. Proteomics* 4, 1134–1144
- Cox, D. M., Zhong, F., Du, M., Duchoslav, E., Sakuma, T., and McDermott, J. C. (2005) Multiple reaction monitoring as a method for identifying protein posttranslational modifications. *J. Biomol. Tech.* **16**, 83–90
- Beynon, R. J., Doherty, M. K., Pratt, J. M., and Gaskell, S. J. (2005) Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nat. Methods* 2, 587–589
- 28. Touitou, Y., Touitou, C., Bogdan, A., Reinberg, A., Auzeby, A., Beck, H., and

Guillet, P. (1986) Differences between young and elderly subjects in seasonal and circadian variations of total plasma proteins and blood volume as reflected by hemoglobin, hematocrit, and erythrocyte counts. *Clin. Chem.* **32**, 801–804

- Atalar, E., Ozmen, F., Haznedaroglu, I., Acil, T., Ozer, N., Ovunc, K., Aksoyek, S., and Kes, S. (2002) Effects of short-term atorvastatin treatment on global fibrinolytic capacity, and sL-selectin and sFas levels in hyperlipidemic patients with coronary artery disease. *Int. J. Cardiol.* 84, 227–231
- van Helden, W. C., Kok-Verspuy, A., Harff, G. A., and van Kamp, G. J. (1985) Rate-nephelometric determination of fibronectin in plasma. *Clin. Chem.* **31**, 1182–1184
- 31. Deleted in proof
- Lestin, M., Hergert, M., Lestin, H. G., Brinker, C., Storm, H., Kuhrt, E., Kuhrt, B., Lambrecht, H. G., and Kuhnel, W. (2002) Evaluation of the chemiluminescence immunoassays for the measurement of troponin I, myoglobin and CK-MB using the IMMULITE system in comparison to other measuring systems. *Clin. Lab.* 48, 211–221